ASPEN PUBLISHERS

The COMPUTER & INTERNET Lawyer

Volume 26 ▲ Number 7 ▲ JULY 2009

Ronald L. Johnston, Arnold & Porter, LLP Editor-in-Chief*

Bots, Scrapers, and Other Unwanted Visitors to Your Web Site: Can You Keep Them Out?

By John F. Zabriskie

Your company's Web site is an invitation to anyone with Internet access to visit it. Of course, you want customers and potential customers to be able to find your site and to easily use it. But people interested in doing business with you probably are not the only visitors to your site. In addition, automated visitors—known as bots, spiders, and crawlers—regularly access your company's site. Some of these automated visitors are benign, and you likely will welcome them as contributors to the success of your site (and your company). For example, search engines periodically crawl sites to gather information to add to the databases that they consult in responding to search requests. Most site owners will choose to permit crawling by Google and Yahoo as crucial means to increasing site traffic.

Not all automated visitors, however, may be so welcome. Some access Web sites to "scrape" information from it for people who have no intention of becoming customers. In several recent and well-publicized disputes, media organizations have complained that other Web sites are aggregating news headlines and lead sentences of stories and that bloggers are using extensive excerpts of the original stories. Both practices may result in the aggregating site or the blog post ranking higher in Google search results than the original story.

But concern over scraping should not be limited to online publishers. Virtually all companies with Web sites, particularly sites allowing access to large product catalogs or other databases, need to be aware of the potential danger posed by scraping. Scraping can overwhelm a Web site's server with an avalanche of automated queries, which in turn will impair the site's usability for intended users. Scraping also could allow a competitor to gain access to your company's information, such as pricing or a directory of articles intended to be

John F. Zabriskie is a partner with Foley & Lardner LLP and a member of the firm's General Commercial Litigation and Intellectual Property Litigation Practices. He has 20 years' experience litigating software, and commercial matters. He has tried cases, both to juries and the bench, in federal and state courts around the country. He has been identified by *The Legal 500 US* as a leading trade secrets lawyer and has represented clients in matters involving virtually all aspects of computer software.

Wolters Kluwer
Law & Business

available to subscribers. Recently, a securities research firm named Newriver, Inc., sued financial information giant Morningstar for allegedly scraping content from Newriver's proprietary online database of prospectus summaries. Even Google and other major search engines are developing new techniques for searching the currently non-crawlable, so-called "Deep Web," that is, Internet-accessible databases that at present are essentially unseen by search engines.

Can your company stop unwanted scraping of its Web site, while maintaining the site's public accessibility and usability? As discussed below, because technological fixes only go so far, you likely also will need to resort to legal solutions to help you strike the right balance. Before discussing those solutions, however, a basic understanding of the technical underpinnings of scraping will be helpful.

How Does Scraping Work?

Scraping is done by software applications that perform automated tasks against Web sites. Applications that have defined Web site targets are called bots. Specialized bots that search multiple Web sites and Web pages are called spiders or crawlers. Bots mimic human users, but perform their designated tasks far faster than a person can. For example, crawlers are essential to Google and other major search engines. To be able to respond swiftly and accurately to search queries, these search engines send out bots to the Internet's more than one trillion Web pages. The bots then download the pages to the search engines' computers, index their content, store the results of the indexing and a copy of the downloaded pages, and follow the links on those pages to yet other pages. This process then repeats itself essentially endlessly. When users enter search terms, the engine checks those terms against its catalogue to generate search results. Because of the enhanced visibility that a Web site gains from being indexed by search engines, virtually all Web site owners welcome the presence of search engine spiders.

Scraping software generally works the same way as search engine Web crawlers. A competitor directs the bot to the targeted site. The bot then downloads the site's pages, scans (or parses) it for the relevant data such as pricing, and stores the retrieved data for use by the competitor.

Stopping Scraping

How does a company prevent unwanted bots from visiting its site or restricted portions of the site? Obviously, there are technical solutions, but none is completely effective against a creative and determined bot designer. For example, while bots sent from the IP address of an offender can be blocked, it is relatively easy for an offender to mask the address or to use an alternative one. Similarly, a Web site can limit access beyond the homepage by requiring users to verify that they are real people by retyping CAPTCHAS, which are distorted characters that bots have difficulty reading.[1] Unfortunately, it is difficult and annoying for people to navigate through CAPTCHAS, so using them may limit use of your Web site by desired visitors. Yet another approach is to use a robot exclusion protocol, which is a file stored in the Web site's server. This file—called robots.txt—contains instructions for bots about the site, such as an instruction not to visit any pages or certain pages on this site. These instructions, however, are not a technical lock, and compliance with them is voluntary. Thus, the benign bots sent from large search engines likely will abide by the instructions, while less well-intended bots are written to sidestep them.

To supplement technical deterrents, there are several legal grounds available to companies to stop unwanted scraping. The most viable ones are discussed below.

Enforcement of Web Site Terms of Use

The first and perhaps most important line of defense are your site's terms of use. Terms of use typically condition use of the site and access to any of its pages or functionality on compliance with the stated terms. Terms of use may be displayed in full or, more typically, accessible through a hyperlink on the site's homepage and other pages. Having terms of use is critical because they are the foundation for most of the available legal remedies.[2]

Early decisions cast some doubt on whether Web site terms of use that did not require explicit agreement—such as the clicking of "I Agree" buttons that are the hallmark of clickwrap licenses—created enforceable contracts between the user and Web site sponsor.[3] More recently, however, courts regularly enforce browsewrap agreements based on such terms so long as the terms, or hyperlinks to the terms, are displayed conspicuously.[4]

These more recent decisions enforcing browsewrap agreements certainly are on solid ground when it is a real person accessing the site because the law commonly imputes to people knowledge of form contract terms.[5] But can there be a manifestation of assent sufficient to create an enforceable contract when the Web site is accessed by an electronic agent such as a bot that cannot understand words and that does not report on either the presence or content of the terms of use? Although the matter is not entirely free from doubt, most courts appear willing to find that bots can enter into enforceable browsewrap agreements on behalf of the humans directing them, at least when bots frequently access a site.[6] In essence, these courts are saying that repeated enjoyment of another person's site will be deemed to

constitute consent to the conditions that the other person imposes for access to that site, regardless of actual notice of the conditions.

These decisions embrace the common sense policy that a person should not be allowed to dodge otherwise binding terms of use by resort to technological sleight of hand. Perhaps acknowledging that more than policy is needed to satisfy the fundamental contractual requirement of assent, however, many of these courts look to decidedly non-digital sources for evidence of that assent—a cease-and-desist letter from plaintiff or its counsel or an admission of actual knowledge by a human representative of the defendant.[7] Thus, without some evidence of actual knowledge by the person directing the bots, this line of cases may be less persuasive.

At the end of the day, however, the fact that some courts are willing to find assent by non-consumer/competitors to terms of use without actual knowledge should be a sign to scrapers that courts will look closely at their behavior. For the same reason, companies are well-advised to review their sites' terms of use to ensure that they prohibit expressly and specifically the type of access, such as scraping by bots or retrieving or copying information for commercial purposes, about which they are most worried. For example, a clause stating that "You agree that you will not use any robot, spider or other automated device, process, or means to access the Site" has been held, unremarkably, to bar use of bots to access a site.[8]

Computer Fraud and Abuse Act

The federal Computer Fraud and Abuse Act (CFAA)[9] was enacted as a criminal statute to prevent outside hacking of government computers. Congress later amended it to cover computers "used in interstate commerce or communication" (which includes virtually all computers connected to the Internet, including servers hosting Web sites) and to provide a civil remedy to "any person who suffers damage or loss by reason of a violation" of the CFAA.[10] These amendments have made CFAA claims ubiquitous in civil litigation, particularly in suits by employers alleging that former employees stole proprietary information from the employer's computers.

The CFAA also is a potential weapon against competitors that scrape or otherwise misuse your company's Web site, even if your company does not own the server hosting the site.[11] Compared to a contract claim for breach of terms of use, a CFAA claim provides a sure basis for federal court jurisdiction and a potentially easier avenue to preliminary relief.

In the context of Web site scraping, two elements of a CFAA civil claim may be problematic to satisfy. First, virtually all sections of the CFAA that might be useful to a private civil plaintiff require that the defendant either "exceeds authorized access" to a protected computer or accesses a protected computer "without authorization."[12] Courts disagree over what it takes to satisfy that language, particularly in disloyal employee cases. One line of authority holds that a person who has access rights to a computer nonetheless either acts without authorization or exceeds that authorization when the computer is accessed with the intent of using the retrieved information for an improper purpose, such as to compete, or when the person accessing has acquired an adverse interest.[13] The opposing line of authority holds that the CFAA addresses only unauthorized access, not authorized access for an improper purpose.[14]

These decisions embrace the common sense policy that a person should not be allowed to dodge otherwise binding terms of use by resort to technological sleight of hand.

While CFAA case law involving Web sites is more settled on this issue, it is not unanimous. Many courts readily and with little analysis conclude that visits by a bot to a competitor's Web site are without access or exceed authorized access when the visit violates the site's terms of use.[15]

At least one recent opinion, however, has expressed skepticism with that conclusion.[16] In that case, the defendant's business was devoted to obtaining advance, preferred boarding passes for Southwest ticket holders by accessing the Southwest Web site on behalf of the ticket holders. Southwest Airlines complained that its site's terms of use expressly prohibited that type of use. While the court had no trouble concluding that the defendant had violated the terms of use, the court nevertheless denied Southwest's motion for summary judgment on its CFAA claim. Noting that the defendant was using the Southwest site consistent with its intended function and was not an outside hacker, the court expressed its reservation about having application of the CFAA—a criminal statute—turn on individual contracts and their interpretation.[17]

The second controversial aspect of asserting a CFAA claim for scraping a Web site involves whether there has been "damage or loss" under the statute. To be able to pursue a civil action under the CFAA, a plaintiff must allege "damage or loss" due to a violation of the CFAA, and (except in highly exceptional cases) a "loss" in excess of $5,000 in one year.[18] "Damage" is defined as "any impairment to the integrity or availability of data,

a program, a system, or information,"[19] while "loss" is defined as "any reasonable cost to any victim, including the cost of responding to an offense, conducting a damage assessment, and restoring the data, program, system, or information to its condition prior to the offense, and any revenue lost, cost incurred, or other consequential damages incurred because of interruption of service."[20]

Establishing damage due to a scraping of a Web site is relatively straightforward when the number and frequency of bot visits clog the site's server, slowing the site's performance or even its availability to the public. That circumstance falls squarely within the term's definition of involving "impairment to the . . . availability of data . . . or information." The courts are split, however, as to whether damage can be shown when there is no such impairment and instead the only consequence is that a competitor has comprehensive knowledge of your company's information (such as prices for thousands of products).[21]

It also is uncomplicated to establish loss under the CFAA when a scraping episode requires spending more than $5,000 to investigate and remedy any damage to the server hosting the site. When the remedial costs do not relate directly to a computer, however, satisfying the loss requirement is less certain.[22] If the loss is due to misuse of information after it is taken from the site, satisfying the loss requirement is even more difficult because the CFAA seems to limit clearly the definition of loss to revenue lost "because of interruption of service."[23] Nonetheless, some courts have held that economic damages unrelated to any interruption of service will qualify as loss.[24]

Copyright Infringement

Scrapers and other uninvited visitors to your company's Web site also may run afoul of the US Copyright Act.[25] Whether the Copyright Act applies to scraping involves several issues, and as is often true in technology cases, resolution of them will depend in large part on how the technology in question actually works. The following discussion stems from the earlier discussion of the technical underpinnings of scraping.

The first question is whether a bot copies anything on the Web site that is protected by the Copyright Act. Generally, copyright protection extends only to original expressions fixed in a tangible medium and does not reach facts or ideas.[26] For Web sites, copyright protection will extend to original, expressive content and the underlying code that produces a Web site display. It also may (or may not!) cover such things as a site's look and feel, but it almost certainly will not cover prices or product codes.[27] Thus, a competitor's retaining a copy of all product prices scraped from a site probably will not, in and of itself, constitute copyright infringement.

But was there a copyright violation arising from the manner in which the scraping program obtained those facts in the first place? Bots (and the humans directing them) do not view Web sites. Instead, as noted, bots generally retrieve from the Web site's server all of the code comprising the site (and its constituent pages) and then store the code in the random access memory (RAM) of the machine running the bot. The bot then rapidly (within seconds) extracts the target data (such as prices) and flushes the rest from the machine's RAM. The brief existence in RAM of a copy of the site code, however, generally will be deemed to be sufficient to have resulted in the creation of a copy for copyright purposes.[28]

Your site's terms of use, assuming that they prohibit bots, may be your trump card, even in the face of an otherwise viable fair use defense.

Is that copy infringing? Here again your site's terms of use are critical. If the terms of use prohibit copying by a bot or copying for a commercial purpose, then there will be infringement because using a protected work outside the scope of a license constitutes copyright infringement.[29]

Demonstrating infringement, however, may not be enough if the scraper has a viable defense. Most likely, the scraper will assert the statutory defense of fair use.[30] In doing so, a scraper will attempt to analogize to software reverse engineering cases such as *Sega Enterprises Ltd. v. Accolade, Inc.*,[31] which hold that the "intermediate copying" of a software program resulting from running the program on the computer of the reverse engineers is a "fair use" if it is the only way to access the unprotected aspects of the code and the copier has a legitimate reason for seeking that access. In the context of Web sites, scrapers have argued with success that retrieving an entire copy of a Web page (including protectable and non-protectable elements) and storing it in a computer's RAM or cache is a fair use when the purpose is to allow a bot to extract non-protectable matter such as prices.[32]

A fair use finding, however, is not inevitable. For example, the reverse engineering cases limit permissible copying to only what was necessary to get at unprotectable matter.[33] Arguably, scraping never is necessary because theoretically a person could review the targeted Web site and manually record it.[34] That argument may

be vulnerable to claims of impracticability and the technological fact that caches only keep as much as necessary to easily reload a page, however.[35]

Ultimately, your site's terms of use, assuming that they prohibit bots, may be your trump card, even in the face of an otherwise viable fair use defense. Despite the fact that fair use is a statutory defense, courts generally enforce contractual agreements that waive the defense or prohibit the conduct on which such a defense is based.[36]

Digital Millennium Copyright Act

Scraping also may violate the federal Digital Millennium Copyright Act (DMCA).[37] Among other things, the DMCA prohibits the circumvention of technological measures that effectively control access to a work protected by a US copyright,[38] as well as the trafficking in devices designed to circumvent such measures. Although a Web site's terms of use establish the conditions for accessing a site, they are contract terms and almost certainly will not qualify as a technological measure. Accordingly, violating those terms probably will not also violate the DMCA. CAPTCHAs—those distorted characters that some sites require users to retype to verify that the user is a person, not a bot—have been held to be a technological measure under the DMCA, however.[39]

The robots.txt file also is intended to keep uninvited bots away from a site. As noted, however, compliance with the instructions in such a file is voluntary. For this reason, the robot exclusion protocol as a general matter probably does not qualify as a technological measure under the DMCA.[40] Accordingly, instructing a bot to ignore a robots.txt file that is part of an active Web site probably will not constitute a violation of the DMCA.[41]

Trespass to Chattels

A Web site owner also may be able to use the common law tort of trespass to chattels against a scraper. Early in the Internet age, some courts began applying this ancient tort to defendants sending spam email to Internet service providers on the basis of assertions that a large volume of spam endangered the functioning of the provider's servers.[42] Online auction site eBay then successfully asserted this claim under California law against an online auction aggregator that used bots to crawl eBay's site for auction information.[43] Notably, the court found that eBay had a likelihood of success on this theory even though the impact from the defendant's bots (approximately 1 percent to 2 percent of eBay's daily queries and data transferred) in and of themselves used "only a small amount of eBay's computer system capacity."[44] Despite the lack of substantial interference with eBay's servers and perhaps reflective of the case arising in the context of a preliminary injunction, the *eBay* court nonetheless was concerned with the threat that eBay would suffer irreparable injury if crawling by other aggregators rendered eBay's site inaccessible to users.[45] The following year, another judge in that same district held that mere use of the plaintiff's computer was sufficient to establish a trespass to chattels claim, even if that use placed only a "negligible" load on the computer.[46]

In 2003, however, the California Supreme Court reined in the potentially broad reach of this tort in *Intel Corp. v. Hamidi*.[47] In that case, which arose in the context of a former Intel employee's bombarding Intel's email system with spam, the California Supreme Court held that this claim could be applied to electronic communication only when that communication either damaged the email system or impaired its functioning.[48]

In light of *Hamidi*, at least in California, asserting this tort against a scraper will not succeed without proof that the scraping is causing some damage to the server hosting the Web site or degrading its usability by inundating it with queries.[49]

Outside California, however, such a claim may be more successful.[50]

Increasing Your Chances of Legal Success

The use of bots and scrapers for competitive advantage is likely to increase. As a result, businesses need to be alert for the presence of these automated and potentially pernicious visitors to their Web sites. The success of any of the legal remedies discussed in this article against unwanted bots and scrapers in any particular case will depend, of course, on the case's unique facts. Even before litigation looms, however, your company can take several steps to enhance its likelihood of prevailing against the sender of the offending bots. These steps include:

- Making sure that at least your site's homepage—but preferably every interior page as well—displays conspicuously at least a link to the terms of use;

- Confirming that your terms of use bar the behavior that you are worried about. Take special care to include an explicit prohibition of the behavior about which you are most concerned. For example, Southwest Airlines included language in its site's terms of use that prohibited third parties from using the site for commercial purposes and from obtaining boarding passes for passengers. When a third party started a

business charging passengers to have a bot get boarding passes from the site, Southwest Airlines succeeded in enforcing that express prohibition against the third party.

- Using the robot exclusion protocol. Even though some bots may ignore a robot.txt file, having that file in place will be helpful in establishing that the bots' presence is unauthorized. By the same token, failure to use this industry-standard protocol may be deemed as implied consent to the bots visiting your site.

- Monitoring the logs of your site's server for evidence of unusual activity that might reveal the existence of unwanted visitors;

- If you detect the presence of unwanted bots, identify the sender, block the IP address used by the sender, and serve a cease-and-desist letter that gives the sender actual notice of the terms of use; and

- If the cease-and-desist letter is not successful, consider seeking preliminary injunctive relief under one or more of the theories discussed in this article.

Notes

1. CAPTCHA is an acronym for Completely Automated Public Turing test to tell Computers and Humans Apart.
2. *See* EF Cultural Travel BV v. Zefer Corp., 318 F. 3d 58, 63-64 (1st Cir. 2003) ("[W]e think that the public Web site provider can easily spell out explicitly what is forbidden . . . . If [a Web site provider] wants to ban scrapers, let it say so on the Web page or a link clearly marked as containing restrictions. . . [P]ublic Web site providers ought to say just what non-password protected access they purport to forbid.").
3. *See, e.g.*, Specht v. Netscape Communications Corp., 306 F.3d 17 (2d Cir. 2002) (users downloading software from Netscape's Web site not bound by arbitration agreement in site's terms of use because those terms were not adequately displayed).
4. *See, e.g.*, Southwest Airlines Co. v. Boardfirst, L.L.C, No. 3:06-CV-0891-B, 2007 WL 4823761 (N.D. Tex. Sept. 12, 2007).
5. *See, e.g.*, Ticketmaster Corp. v. Tickets.com, Inc., No. CV997654HLHVBKX, (C.D. Cal. Mar. 7, 2003) (in enforcing Web site terms of use, court deemed those terms analogous to terms printed on the back of airline and parking lot tickets).
6. *See, e.g.*, Cairo, Inc. v. Crossmedia Services, Inc., No. C 04-04825 JW, 2005 WL 756610 (N.D. Cal. Apr. 1, 2005) (holding that defendant's "repeated and automated use of [plaintiff's] web pages can form the basis of imputing knowledge to [defendant] of the terms" of use); *see also* Register.com, Inc. v. Verio, Inc., 356 F.3d 393, 402-403 (2d Cir. 2004) (same).
7. *See, e.g.,* Ticketmaster L.L.C. v. RMG Technologies, Inc., 507 F. Supp. 2d 1096, 1107 (C.D. Cal. 2007) (defendant conceded being on notice of terms of use); Southwest v. Boardfirst, 2007 WL 4823761 *7 (the crawling found to be improper occurred after defendant received cease and desist letter); *Register.com, Inc.,* 356 F.3d at 402 (defendant admitted awareness of terms of use); *Cairo,* 2005 WL 756610 *4 (same); *but see* Internet Archive v. Shell, 505 F. Supp. 2d 755 (D. Colo. 2007) (denying motion to dismiss counterclaim by Web site owner that non-profit Web site archive breached site's terms of use where site accessed only by archive's spiders and there was no allegation that archive had actual knowledge of the terms of use).
8. *See Ticketmaster, L.L.C.*, 507 F. Supp. 2d at 1107-1109.
9. 18 U.S.C. § 1030.
10. *Id.*, §§ 1030(e)(2)(B) and (g).
11. *See, e.g.,* Theofel v. Farey-Jones, 359 F.3d 1066, 1078 (9th Cir. 2004) ("Individuals other than the computer's owner may be proximately harmed by unauthorized access, particularly if they have rights to data stored on it.").
12. *See, e.g.,* 18 U.S.C. §§ 1030(a)(2), (4), and (5).
13. *See, e.g.*, Int'l Airports Center L.L.C. v. Citrin, 440 F.3d 418 (7th Cir. 2006).
14. *See, e.g.,* Shamrock Food Co. v. Gast, 535 F. Supp. 2d 962 (D. Ariz. 2008).
15. *See, e.g.,* Southwest Airlines Co. v. FareChase, Inc., 318 F. Supp. 2d 435, 439-340 (N.D. Tex. 2004); Register.com v. Verio, Inc., 126 F. Supp. 2d 238, 251-253 (S.D.N.Y. 2000), *aff'd on other grounds,* 356 F.3d 393 (2d Cir. 2004); *Ticketmaster, L.L.C.*, 507 F. Supp. 2d at 1113.
16. Southwest Airlines Co. v. Boardfirst, 2007 WL 4823761.
17. *Id.* at *12-14. This same concern entered the public spotlight with the 2008 prosecution of Lori Drew under the CFAA. In that case, Drew, an adult, was convicted of misdemeanor charges for violating the terms of use of the social networking site MySpace. Drew registered under a false name as a child and then harassed another child through MySpace posts, conduct that tragically ended in the other child's committing suicide. Drew is trying to set aside the conviction.
18. 18 U.S.C. §§ 1030(g), (a)(5)(B)(i).
19. *Id.* § 1030(e)(8).
20. *Id.* § 1030(e)(11).
21. *Compare* Therapeutic Research Faculty v. NBTY, Inc., 488 F. Supp. 2d 991, 996 (E.D. Cal. 2007) (unauthorized access to online subscription database may constitute impairment to data and thus fall within the CFAA's definition of "damage," even though data was not changed or erased) *with* Garelli Wong & Assocs., Inc. v. Nichols, 551 F. Supp. 2d 704 (N.D. Ill. 2008) (defendant's copying of employer's confidential information alone insufficient to constitute "damage" under CFAA).
22. *Compare* Patrick Patterson Custom Homes, Inc. v. Bach, 586 F. Supp. 2d 1026, 1036-1037 (N.D. Ill. 2008) (in case involving employee's transfer of employer's funds to herself and deletion of financial records from employer's laptop, costs of ordering duplicate financial records and of hiring an accountant to

reconstruct those records satisfied CFAA "loss" requirement) *with* Nexans Wires S.A. v. Sark-USA, Inc., 319 F. Supp. 2d 468, 474-476 (S.D.N.Y. 2004) (costs of assessing and remedying business consequences from data being improperly accessed did not relate to a computer and thus did not qualify as "loss" under CFAA).

23. 18 U.S.C. § 1030(e)(11). Even though lost profits may not satisfy the CFAA "loss" requirement, they are still recoverable as "economic damages" in a civil action (assuming that the "loss" requirement otherwise is satisfied). *Id*. § 1030(g).
24. *Therapeutic Research Faculty,* 488 F. Supp. 2d at 996-997 (cost of user license to Web site portion defendant improperly accessed constituted "loss" under CFAA).
25. 17 U.S.C. §§ 101, *et seq*.
26. Feist Publications, Inc. v. Rural Tel. Serv. Co., Inc., 499 U.S. 340 (1991).
27. QSRSoft, Inc. v. Restaurant Technology, Inc., No. 06 C 2734, 2006 WL 2990432 (N.D. Ill. Oct. 19, 2006) (granting injunction where defendant copied look and feel of plaintiff's Web site); *Feist*, 499 U.S. at 351 ("In no event may copyright extend to the facts themselves).
28. MAI Systems Corp. v. Peak Computer, Inc., 991 F.2d 511, 519 (9th Cir. 1993).
29. *Ticketmaster, L.L.C.,* 507 F. Supp. 2d at 1106-1107.
30. 17 U.S.C. § 107.
31. Sega Enterprises Ltd. v. Accolade, Inc., 977 F.2d 1510, 1528 (9th Cir. 1992).
32. Ticketmaster v. Tickets.com, 2003 WL 21406289 at *4.
33. *Sega*, 977 F.2d at 1528.
34. Ticketmaster v. Tickets.com, 2003 WL 21406289 at *4.
35. Perfect 10, Inc. v. Amazon.com, Inc., 487 F.3d 701, 726 (9th Cir.), *amended by* 508 F.3d 1146 (9th Cir. 2007).
36. Bowers v. Baystate Technologies, Inc., 320 F.3d 1317 (Fed. Cir. 2003) (holding that a contractual prohibition of reverse engineering was not pre-empted by the Copyright Act).
37. 17 U.S.C. §§ 1201, *et seq*.
38. As noted, there is some question whether the material that a competitor's bots might copy from your company's Web site, such as pricing, is copyrightable.
39. *See*, *e.g.*, Ticketmaster, L.L.C., 507 F. Supp. 2d at 1112.
40. *See* Healthcare Advocates, Inc. v. Harding, Earley, Follmer & Frailey, 497 F. Supp. 2d 627, 643 (E.D. Pa. 2007) (stating in *dicta* that the robot exclusion protocol "by itself is not analogous to digital password protection or encryption").
41. *Id*. (court qualified its holding that a robots.txt file on a Web site stored in the Internet Archive's Wayback Machine constituted a DMCA "technological measure" by stating that its holding "should not be interpreted as a finding that a robots.txt file universally qualifies as a technological measure" under the DMCA).
42. *See, e.g.*, America Online, Inc. v. IMS, 24 F. Supp. 2d 548 (E.D. Va. 1998).
43. eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058 (N.D. Cal. 2000).
44. *Id*. at 1071.
45. *Id*. at 1071-1072.
46. Oyster Software, Inc. v. Forms Processing, Inc., No. C-00-0724 JCS, 2001 WL 1736382, at *13 (N.D. Cal. Dec. 6, 2001).
47. Intel Corp. v. Hamidi, 71 P.3d 296 (Cal. 2003).
48. *Id*. at 306-307.
49. *See, e.g.*, Ticketmaster Corp. v. Tickets.com, 2003 WL 21406289, *3 (rejecting Web site operator's attempt to assert trespass to chattels claim against scraper where there was no evidence that the bots adversely affected the site's performance).
50. *See., e.g.*, *Register.com, Inc.*, 356 F.3d at 404 (trespass to chattels claim would lie even though defendant's bots "alone would not incapacitate" plaintiff's systems); Southwest Airlines Co. v. Farechase, Inc., 318 F. Supp. 2d at 442 (wrongful interference without damage to or deprivation of use of server sufficient to state trespass to chattels claim). Regardless of whether a plaintiff is required to prove and can prove damage, a trespass to chattels claim may be subject to preemption under the Copyright Act. *See, e.g.*, *Healthcare Advocates, Inc.*, 497 F. Supp. 2d at 649-650.

Reprinted from *The Computer & Internet Lawyer*, July 2009, Volume 26, Number 7, pages 5 to 11, with permission from Aspen Publishers, Inc., a Wolters Kluwer business, New York, NY, 1-800-638-8437, *www.aspenpublishers.com*.